CLAIMS

1. A method for translating a report generated in natural language into structured computer-understandable frames comprising:

eliciting directed input as to a medical condition and symptoms;

using the directed input elicited to identify a disease signature corresponding to the medical condition and symptoms;

using the disease signature to identify a lexical domain containing language information pertinent to the disease signature, the lexical domain having been programmed with word properties for words expected to be used with regard to the disease signature, the word properties including a likelihood that combinations of words in the lexical domain interdepend and each word's inherent tendency to link with other words;

looking up the word properties for words used in the report in the lexical domain;

calculating for combinations of words used in sentences contained in the report a statistical likelihood that they interdepend and identifying probable word links;

semantically interpreting a nature of the probable word links; and

generating the structured computer-understandable frames based on the nature of the probable word links.

2. The method of claim 9 further comprising structurally analyzing the report to delineate between the sections of the report and between separate sentences.

3. The method of claim 9 wherein the word properties include the syntactic properties of each word.

4. The method of claim 9 wherein the word properties include the semantic properties of each word.

5. The method of claim 9 wherein the word properties include both the syntactic and semantic properties of each word.

6. The method of claim 9 wherein the likelihood of that combinations of words in the lexical domain interdepend is based on empirical statistics on how frequently the combinations of words in the lexical domain have interdepended in preexisting reports.

7. The method of claim 9 wherein the inherent tendency of each word to link with other words is based on empirical statistics of how frequently that word interacts with other words and with what types of other words.

8. The method of claim 9 wherein, for novel combinations of individual words not stored in the lexical domain, the likelihood that the novel word combination interdepends is estimated based on syntactic and semantic properties of the individual words.

9. A method for translating a report about a patient, afflicted with a medical condition and symptoms, generated in natural language into structured computer-understandable frames comprising:

identifying a disease signature corresponding to the medical condition and symptoms;

using the disease signature to identify a lexical domain containing language information pertinent to the disease signature, the lexical domain having been programmed with word properties for words expected to be used with regard to the disease signature, the word properties including a likelihood that combinations of words in the lexical domain interdepend and each word's inherent tendency to link with other words;

looking up the word properties for words used in the report in the lexical domain;

calculating for combinations of words used in sentences contained in the report a statistical likelihood that they interdepend and identifying probable word links;

semantically interpreting a nature of the probable word links; and

generating the structured computer-understandable frames based on the nature of the probable word links.

10. The method of claim 9 further comprising structurally analyzing the report to delineate between the sections of the report and between separate sentences.

11. The method of claim 9 wherein the word properties include the syntactic properties of each word.

12. The method of claim 9 wherein the word properties include the semantic properties of each word.

13. The method of claim 9 wherein the word properties include both the syntactic and semantic properties of each word.

14. The method of claim 9 wherein the likelihood of that combinations of words in the lexical domain interdepend is based on empirical statistics on how frequently the combinations of words in the lexical domain have interdepended in preexisting reports.

15. The method of claim 9 wherein the inherent tendency of each word to link with other words is based on empirical statistics of how frequently that word interacts with other words and with what types of other words.

16. The method of claim 9 wherein, for novel combinations of individual words not stored in the lexical domain, the likelihood that the novel word combination interdepends is estimated based on syntactic and semantic properties of the individual words.

17. A system for translating a report generated in natural language into structured computer-understandable frames comprising:

a patient input module that elicits from a patient directed input as to the patient's medical condition and symptoms and, based on the patient directed input, identifies a disease signature corresponding the patient's medical condition and symptoms;

a lexical analyzer using a lexical domain containing language information pertinent to the disease signature, the lexical domain having been programmed with word properties for words expected to be used with regard to the disease signature, the word properties including a likelihood that combinations of words in the lexical domain interdepend and each word's inherent tendency to link with other words, and the lexical analyzer looks up the word properties for words used in the report in the lexical domain;

a parser/semantic interpreter module for calculating for combinations of words used in sentences contained in the report a statistical likelihood that they interdepend and identifying probable word links; and

a structured frame generator that creates the structured computer-understandable frames based on the nature of the probable word links.

18. The system of claim 17 further comprising a structural analyzer to delineate the report into sections and delineate between separate sentences.

19. The system of claim 17 wherein the word properties include the syntactic properties of each word.

20. The system of claim 17 wherein the word properties include the semantic properties of each word.

21. The system of claim 17 wherein the word properties include both the syntactic and semantic properties of each word.

22. The system of claim 17 wherein the likelihood of that combinations of words in the lexical domain interdepend is based on empirical statistics on how frequently the combinations of words in the lexical domain have interdepended in preexisting reports.

23. The system of claim 17 wherein the inherent tendency of each word to link with other words is based on empirical statistics of how frequently that word interacts with other words and with what types of other words.

24. The system of claim 17 wherein, for novel combinations of individual words not stored in the lexical domain, the likelihood that the novel word combination interdepends is estimated based on syntactic and semantic properties of the individual words.

25. A system for translating a report about a patient, afflicted with a medical condition and symptoms, generated in natural language into structured computer-understandable frames comprising:

a disease signature identifier which identifies a disease from which the patient is suffering corresponding to the patient's medical condition and symptoms;

a lexical analyzer using a lexical domain containing language information pertinent to the disease signature, the lexical domain having been programmed with word properties for words expected to be used with regard to the disease signature, and the lexical analyzer looks up the word properties for words used in the report in the lexical domain;

a parser/semantic interpreter for calculating for combinations of words used in sentences contained in the report a statistical likelihood that they interdepend and identifying probable word links; and

a structured frame generator that creates the structured computer-understandable frames based on the nature of the probable word links.

26. The system of claim 25 further comprising a structural analyzer to delineate the report into sections and delineate between separate sentences.

27. The system of claim 25 wherein the word properties include the syntactic properties of each word.

28. The system of claim 25 wherein the word properties include the semantic properties of each word.

29. The system of claim 25 wherein the word properties include both the syntactic and semantic properties of each word.

30. The system of claim 25 wherein the likelihood of that combinations of words in the lexical domain interdepend is based on empirical statistics on how frequently the combinations of words in the lexical domain have interdepended in preexisting reports.

31. The system of claim 25 wherein the inherent tendency of each word to link with other words is based on empirical statistics of how frequently that word interacts with other words and with what types of other words.

32. The system of claim 25 wherein, for novel combinations of individual words not stored in the lexical domain, the likelihood that the novel word combination interdepends is estimated based on syntactic and semantic properties of the individual words.